

# SHARP ESTIMATES FOR THE CONVERGENCE RATE OF ORTHOMIN(K) FOR A CLASS OF LINEAR SYSTEMS

ANDREI DRĂGĂNESCU\* AND FLORIN SPINU†

**Abstract.** In this work we show that the convergence rate of Orthomin( $k$ ) applied to systems of the form  $(I + \rho U)x = b$ , where  $U$  is a unitary operator and  $0 < \rho < 1$ , is less than or equal to  $\rho$ . Moreover, we give examples of operators  $U$  and  $\rho > 0$  for which the asymptotic convergence rate of Orthomin( $k$ ) is exactly  $\rho$ , thus showing that the estimate is sharp. While the systems under scrutiny may not be of great interest in themselves, their existence shows that, in general, Orthomin( $k$ ) does not converge faster than Orthomin(1). Furthermore, we give examples of systems for which Orthomin( $k$ ) has the same asymptotic convergence rate as Orthomin(2) for  $k \geq 2$ , but smaller than that of Orthomin(1). The latter systems are related to the numerical solution of certain partial differential equations.

**Key words.** Orthomin( $k$ ), iterative methods

**AMS subject classifications.** 65F10

**1. Introduction.** Developed by Vinsome [8], Orthomin( $k$ ) ( $k = 1, 2, 3, \dots$ ) is a family of iterative methods for solving linear systems of the form

$$Ax = b, \quad (1.1)$$

where  $A \in M_d(\mathbb{C})$  is a nonsingular, possibly non-symmetric matrix, and  $b \in \mathbb{C}^d$ . If  $x_n$  is the  $n^{\text{th}}$  iterate of the method and  $r_n = b - Ax_n$  is the  $n^{\text{th}}$  residual, the idea behind Orthomin( $k$ ) is to compute the next iterate of the form  $x_{n+1} = x_n + x$  with the correction  $x$  residing in a  $k$ -dimensional subspace  $V_k^{(n)}$  (or  $(n+1)$ -dimensional for  $(n+1) < k$ ) and the Euclidean norm of the next residual  $r_{n+1} = r_n - Ax$  minimized:

$$\|r_{n+1}\| = \min_{x \in V_k^{(n)}} \|r_n - Ax\|. \quad (1.2)$$

The definition (1.2) is equivalent to

$$r_{n+1} = r_n - \Pi_{AV_k^{(n)}} r_n, \quad (1.3)$$

where  $\Pi_V$  is the orthogonal projection on a subspace  $V$ . The algorithm generates a sequence of vectors  $p_0, p_1, p_2, \dots$ , called the search directions, and for  $n \geq k-1$  the space  $V_k^{(n)}$  is generated by the last  $k$  search directions  $p_n, p_{n-1}, \dots, p_{n-k+1}$ ; for  $n < k-1$  the space  $V_k^{(n)}$  is simply  $\text{span}\{p_n, p_{n-1}, \dots, p_0\}$ . To give a precise formulation, for an initial guess  $x_0$  we initialize the residual and the initial search direction by  $p_0 = r_0 = b - Ax_0$ , and the Orthomin( $k$ ) iteration reads

$$x_{n+1} = x_n + \lambda_n p_n, \quad r_{n+1} = r_n - \lambda_n A p_n, \quad (1.4)$$

$$p_{n+1} = r_{n+1} - \sum_{j=1}^{\min(k-1, n+1)} \nu_n^{(j)} p_{n-j+1}, \quad (1.5)$$

\*Department of Mathematics and Statistics, University of Maryland, Baltimore County, 1000 Hill-top Circle, Baltimore, Maryland 21250 ([draga@math.umbc.edu](mailto:draga@math.umbc.edu)). The work of this author was supported in part by the Department of Energy under contract no. de-sc0005455, and by the National Science Foundation under awards DMS-1016177 and DMS-0821311.

†Campbell & Co, 2850 Quarry Lake Drive, Baltimore, Maryland 21209 ([fspinu@gmail.com](mailto:fspinu@gmail.com)).

with

$$\lambda_n = \frac{(r_n, Ap_n)}{(Ap_n, Ap_n)}, \quad \nu_n^{(j)} = \frac{(Ar_{n+1}, Ap_{n-j+1})}{(Ap_{n-j+1}, Ap_{n-j+1})}, \quad j = 1, \dots, \min(k-1, n+1), \quad (1.6)$$

where  $(u, v) = \sum_{j=1}^d u_j \overline{v_j}$  denotes the inner product in  $\mathbb{C}^d$ , and  $\|u\| \stackrel{\text{def}}{=} \sqrt{(u, u)}$ . The coefficients  $\lambda_n$  and  $\nu_n^{(j)}$  in (1.6) are defined so that

$$r_{n+1} \perp Ap_n \quad \text{and} \quad Ap_{n+1} \perp Ap_{n-j+1}, \quad j = 1, \dots, \min(k-1, n+1). \quad (1.7)$$

A simple inductive argument shows that  $r_{n+1} \perp Ap_{n-j+1}$  for  $j = 1, \dots, \min(k, n+1)$ , and hence (1.2) holds.

Orthomin( $k$ ) is attractive for a number of reasons. First, as with other iterative methods, Orthomin( $k$ ) can be implemented so that each iteration requires only one matrix-vector (mat-vec) at an additional cost of  $O(kd)$  Flops; a maximum number of  $k$  vectors need to be stored. This is different from GMRES where both the computational cost per iteration and the storage requirements increase with the iteration number. Finally, when symmetric positive preconditioners are used to produce a split preconditioning of Orthomin( $k$ ), the preconditioned iteration can be implemented without reference to the factors of the preconditioners. This is a feature shared with CG, as shown by Elman [3].

In terms of convergence properties, Orthomin( $k$ ) is guaranteed to converge if the field of values<sup>1</sup> of the matrix  $A$  does not contain the origin. The precise convergence result and estimate shown below appears in [4] as Theorem 2.2.2, and was proved first by Eisenstat, Elman, and Schultz in [2] (see also Elman [3]) for matrices with positive definite symmetric part. We recall the result in [4] as

**THEOREM 1.1.** *Assume that  $0 \notin \mathcal{F}(A)$  and  $\delta = \text{dist}(0, \mathcal{F}(A))$ . If  $r_n$  is the  $n^{\text{th}}$  residual in the Orthomin( $k$ ) iteration, then*

$$\|r_{n+1}\| \leq \|r_n\| \sqrt{1 - \frac{\delta^2}{\|A\|^2}}, \quad (1.8)$$

where  $\|A\|$  is the 2-norm of the matrix  $A$ .

We also recall from [4] the parallelism between Orthomin(1) and steepest descent one one hand, and between Orthomin(2) and the conjugate gradient method (CG) on the other. Steepest descent can only be used in connection to symmetric positive definite (SPD) systems and has an iteration of the form (1.4) with the search direction given by  $p_n = r_n$ , just like Orthomin(1). However, for steepest descent the coefficient  $\lambda_n$  is chosen so that

$$e_{n+1} = e_n - \Pi_{\text{span}\{r_n\}}^A e_n,$$

where  $\Pi_V^A$  is the projection on the subspace  $V$  with respect to the  $A$ -inner product  $(u, v)_A = (Au, v)$ . Consequently, the error estimates for steepest descent are similar to the ones for Orthomin(1), and in practice the two methods converge comparably fast for SPD systems. Analogously, the sequence of search directions  $p_0, p_1, \dots$  for CG follows a recursion that is similar to Orthomin(2), except for in CG we have

$$e_{n+1} = e_n - \Pi_{\text{span}\{p_n, p_{n+1}\}}^A e_n.$$

---

<sup>1</sup>The *field of values* or *numerical range* of a complex matrix  $A$  is defined as the set of complex numbers  $\mathcal{F}(A) = \{(Au, u) : u \in \mathbb{C}^d, \|u\| = 1\}$ .

In addition, in the case of CG the second set of orthogonality relations in (1.7) is replaced by the  $A$ -orthogonality relation  $p_{n+1} \perp_A p_n$  (conjugacy), whereas for Orthomin(2) they read  $Ap_{n+1} \perp Ap_n$ . Even though the superiority of CG over steepest descent is well established and understood, not the same can be said about the relation of Orthomin(2) with Orthomin(1) for nonsymmetric systems. We quote from Anne Greenbaum's text [4] (p.34): "Unfortunately, no stronger a priori bounds on the residual norm are known for Orthomin(2) applied to a general matrix whose field of values does not contain the origin although, in practice, it may perform significantly better than Orthomin(1)."

The main contribution of this article is to show that Orthomin( $k$ ) **does not perform better in general** (that is, for matrices  $A$  that satisfy  $0 \notin \mathcal{F}(A)$ ) than Orthomin(1). In Section 2 we consider matrices of the form  $A = I + \rho U$  with  $0 < \rho < 1$  and  $U$  unitary, and we conjecture that, under certain conditions, the asymptotic convergence rate of Orthomin( $k$ ) for such systems is  $\rho$ ; we support our conjecture with numerical evidence for  $k \geq 2$  and we provide analytical arguments for  $k = 1$  in Section 4, which forms the core of this article. Prior to the analysis of the convergence rate of Orthomin(1), in Section 3 we give examples of systems for which Orthomin( $j$ ),  $j = 2, \dots, k$  all achieve the same asymptotic convergence rate, but converge faster than Orthomin(1).

**2. The main examples.** Consider the linear system

$$(I + \rho U)x = b, \quad (2.1)$$

where  $0 < \rho < 1$ ,  $U \in M_d(\mathbb{C})$  is a unitary matrix, and  $b \in \mathbb{C}^d$ . Our goal is to assess the behavior of the ratios

$$q_n = \frac{\|r_{n+1}^{(k)}\|}{\|r_n^{(k)}\|}, \quad (2.2)$$

where  $r_n^{(k)}$  is the  $n^{\text{th}}$  residual in the Orthomin( $k$ ) iteration.

**2.1. An upper bound.** The fact that  $q_n$  is bounded above by  $\rho$  is a consequence of the following result.

**THEOREM 2.1.** *Let  $A \in M_d(\mathbb{C})$  be a normal matrix so that*

$$\sigma(A) \subseteq \overline{\mathcal{B}_\rho(z_0)} \quad (2.3)$$

*with  $0 < \rho < |z_0|$ . The residuals  $r_n^{(k)}$  obtained by applying the Orthomin( $k$ ) iteration to the system (1.1) satisfy*

$$\|r_{n+1}^{(k)}\| \leq \frac{\rho}{|z_0|} \|r_n^{(k)}\|. \quad (2.4)$$

*Proof.* Let  $U = \rho^{-1}(A - z_0 I)$ . Since  $\sigma(A) \subseteq \overline{\mathcal{B}_\rho(z_0)}$  we have  $\sigma(U) \subseteq \overline{\mathcal{B}_1(0)}$ . Because  $A$  is normal it follows that  $U$  is also normal, hence  $\|U\|_2 \leq 1$ . If  $p_0, p_1, \dots$  are the search directions of Orthomin( $k$ ) we have

$$r_{n+1}^{(k)} = r_n^{(k)} - \Pi_{\text{span}\{Ap_n, \dots, Ap_{n-j}\}} r_n^{(k)}$$

where  $j = \min(n, k-1)$ . Hence

$$\|r_{n+1}^{(k)}\| \leq \|r_n^{(k)} - v\|, \quad \forall v \in \text{span}\{Ap_n, \dots, Ap_{n-j}\}.$$

From the construction of the search direction we have  $r_n^{(k)} \in \text{span}\{p_n, \dots, p_{n-j}\}$ , so

$$Ar_n^{(k)} \in \text{span}\{Ap_n, \dots, Ap_{n-j}\} .$$

Therefore

$$\|r_{n+1}^{(k)}\| \leq \|r_n^{(k)} - z_0^{-1}Ar_n^{(k)}\| = \frac{\rho}{|z_0|} \|Ur_n^{(k)}\| \leq \frac{\rho}{|z_0|} \|r_n^{(k)}\| . \quad \square$$

We should note that for normal operators, the field of values being equal to the convex hull of the spectrum [5], condition (2.3) is equivalent to

$$\mathcal{F}(A) \subseteq \overline{\mathcal{B}_\rho(z_0)} .$$

Hence, the general result (1.8) implies

$$\|r_{n+1}^{(k)}\| \leq \sqrt{1 - \frac{(|z_0| - \rho)^2}{(|z_0| + \rho)^2}} \|r_n^{(k)}\| = \frac{2\sqrt{\rho/|z_0|}}{1 + \rho/|z_0|} \|r_n^{(k)}\| . \quad (2.5)$$

The bound (2.4), valid for normal operators only, is significantly sharper than (2.5).

**2.2. Sharpness of the upper bound.** To show that the estimate (2.4) is sharp we consider the diagonal matrices

$$U = \text{diag}[1, \zeta_d, \zeta_d^2, \dots, \zeta_d^{d-1}] , \quad (2.6)$$

where  $\zeta_d = \exp(2\pi i/d)$  is the primitive root of unity of order  $d$ . Without loss of generality we take  $z_0 = 1$ .

**CONJECTURE 2.2.** *For all  $k \in \mathbb{N}$ , there exists  $d_k \in \mathbb{N}$  and  $0 < \rho_k < 1$  so that for all  $\rho \in (0, \rho_k)$  and  $d \geq d_k$ , the residuals  $r_n^{(k)}$  obtained by applying the Orthomin( $k$ ) iteration to the system (2.1) with  $U$  of the form (2.6) and initial value  $x_0 = 0$  satisfy*

$$\lim_{n \rightarrow \infty} \frac{\|r_{n+1}^{(k)}\|}{\|r_n^{(k)}\|} = \rho . \quad (2.7)$$

In this article we prove Conjecture 2.2 for  $k = 1$  (see Theorem 4.14 in Section 4.5). For  $k \geq 2$ , the numerical evidence in support of Conjecture 2.2 is quite strong, as shown in Section 2.3. A consequence of Conjecture 2.2 is that for a given  $k \in \mathbb{N}$  we can find linear systems for which all of Orthomin( $j$ ),  $j = 1, \dots, k$ , achieve the same convergence rate. This shows that, **in general**, Orthomin( $k$ ) does not converge faster than Orthomin(1).

Naturally, for any system in  $\mathbb{C}^d$ , Orthomin( $d$ ) will converge in at most  $d$  steps. In order to find linear systems so that for **all**  $k \in \mathbb{N}$ , Orthomin( $k$ ) achieves the same asymptotic convergence rate as Orthomin(1), we consider the infinite dimensional generalization of (2.6). Let  $d\mu_0(z)$  be the Haar probability measure on the unit circle  $\mathbb{S}^1$ , and consider the operator

$$U : L^2(\mathbb{S}^1, d\mu_0) \rightarrow L^2(\mathbb{S}^1, d\mu_0), \quad Uf(z) = zf(z) , \quad (2.8)$$

and the linear system (2.1) with right-hand side  $b \in L^2(\mathbb{S}^1, d\mu_0)$  given by

$$b(z) = 1, \quad \forall z \in \mathbb{S}^1 .$$

CONJECTURE 2.3. *For all  $0 < \rho < 1$  and  $k \in \mathbb{N}$  the residuals  $r_n^{(k)}$  obtained by applying the Orthomin( $k$ ) iteration to the system (2.1) with  $U$  of the form (2.8),  $b$  as above and zero initial guess satisfy*

$$\lim_{n \rightarrow \infty} \frac{\|r_{n+1}^{(k)}\|}{\|r_n^{(k)}\|} = \rho. \quad (2.9)$$

In Section 4.6 we prove Conjecture 2.3 for  $k = 1$ .

**2.3. Numerical evidence.** In our attempt to verify Conjecture 2.2 we conducted numerical experiments with Orthomin( $k$ ) for the system (2.1) with  $U$  as in (2.6). In Table 2.1 we show the residual norms  $\|r_n^{(k)}\|$  as well as the ratios  $q_n$  for  $d = 13$ ,  $\rho = 0.8$ , and  $k = 1, 2, 3, 4, 5, 7, 9, 11$ . For  $k = 5, 7, 9, 11$  we only show iterates 5–14, as the first 4 iterates are identical to those of Orthomin(4). Note that, in general, the first  $k$  steps of Orthomin( $k+1$ ) are identical to those of Orthomin( $k$ ). The numbers in Table 2.3 show that  $q_n \rightarrow 0.8$  for  $k = 1, 2, 3, 4, 5$ , which is confirmed by further examining the difference  $|r_n - 0.8|$ . However, for  $k = 7, 9, 11$  we notice that  $q_n$  exhibits an oscillatory behavior and does not appear to converge to  $\rho = 0.8$ . This is confirmed by Figure 2.1. Further evidence shows that even for  $k = 7, 9, 11$  the ratio  $q_n$  converges to  $\rho$  if we increase  $d$ . This is confirmed in Figure 2.2. We should point out that we also experimented with random values for  $r_0, b$ , as well as small random perturbations  $\mu_k$  of the numbers  $\zeta_d^k$  (with  $|\mu_k| = 1$ ): we always observe  $q_n \rightarrow \rho$  as long as  $d$  is sufficiently large.

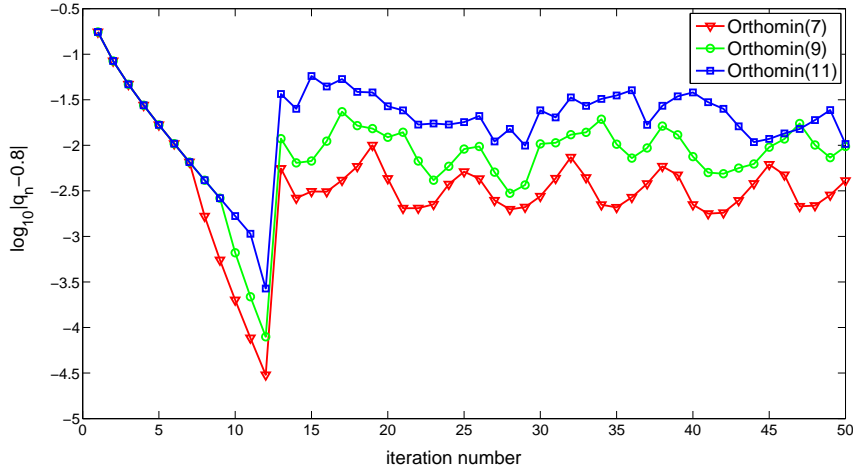


FIG. 2.1.  $\log_{10} |q_n - \rho|$  for Orthomin( $k$ ),  $k = 7, 9, 11$  ( $\rho = 0.8$ ,  $d = 13$ ).

**2.4. Connection with numerical PDEs.** Systems of the form (2.1) arise naturally in the numerical solution of partial differential equations. Consider the steady-state advection-reaction-diffusion equation on  $[0, 2\pi]$

$$-au''(x) + bu'(x) + cu(x) = f(x), \quad a > 0, \quad c \geq 0, \quad b \in \mathbb{R}, \quad (2.10)$$

with periodic boundary conditions  $u(0) = u(2\pi)$ ,  $u'(0) = u'(2\pi)$ . To obtain a discretization of (2.10) we proceed as follows: set  $x_j = jh$ ,  $j = 0, 1, \dots, d$ ,  $h = 2\pi/d$ ,

TABLE 2.1  
Convergence rates for *Orthomin*( $k$ ) for  $d = 13$ ,  $\rho = 4/5$ ,  $k = 1, 2, 3, 4, 5, 7, 9, 11$ .

it	Orthomin(1)		Orthomin(2)		Orthomin(3)		Orthomin(4)	
	$\ r_n\ $	$q_n$	$\ r_n\ $	$q_n$	$\ r_n\ $	$q_n$	$\ r_n\ $	$q_n$
1	3.6056	0.6247	3.6056	0.6247	3.6056	0.6247	3.6056	0.6247
2	2.2524	0.7533	2.2524	0.7156	2.2524	0.7156	2.2524	0.7156
3	1.6967	0.7832	1.6118	0.7768	1.6118	0.7533	1.6118	0.7533
4	1.3289	0.7935	1.2520	0.7919	1.2141	0.7873	1.2141	0.7725
5	1.0544	0.7974	0.9915	0.7958	0.9559	0.7957	0.9379	0.7927
6	0.8407	0.7989	0.7891	0.7985	0.7606	0.7984	0.7435	0.7975
7	0.6717	0.7996	0.6301	0.7993	0.6073	0.7990	0.5929	0.7991
8	0.5371	0.7998	0.5036	0.7997	0.4852	0.7996	0.4738	0.7996
9	0.4296	0.7999	0.4028	0.7999	0.3880	0.7999	0.3789	0.7997
10	0.3436		0.3222		0.3103		0.3030	
it	Orthomin(5)		Orthomin(7)		Orthomin(9)		Orthomin(11)	
	$\ r_n\ $	$q_n$	$\ r_n\ $	$q_n$	$\ r_n\ $	$q_n$	$\ r_n\ $	$q_n$
5	0.9379	0.7832	0.9379	0.7832	0.9379	0.7832	0.9379	0.7832
6	0.7346	0.7956	0.7346	0.7896	0.7346	0.7896	0.7346	0.7896
7	0.5844	0.7985	0.5800	0.7935	0.5800	0.7935	0.5800	0.7935
8	0.4667	0.7995	0.4602	0.7983	0.4602	0.7959	0.4602	0.7959
9	0.3731	0.7998	0.3674	0.7995	0.3663	0.7974	0.3663	0.7974
10	0.2984	0.7999	0.2937	0.7998	0.2920	0.7993	0.2920	0.7983
11	0.2387	0.7999	0.2349	0.7999	0.2334	0.7998	0.2331	0.7989
12	0.1909	0.8000	0.1879	0.8000	0.1867	0.7999	0.1863	0.7997
13	0.1528	0.7999	0.1503	0.7944	0.1493	0.7882	0.1490	0.7634
14	0.1222		0.1194		0.1177		0.1137	

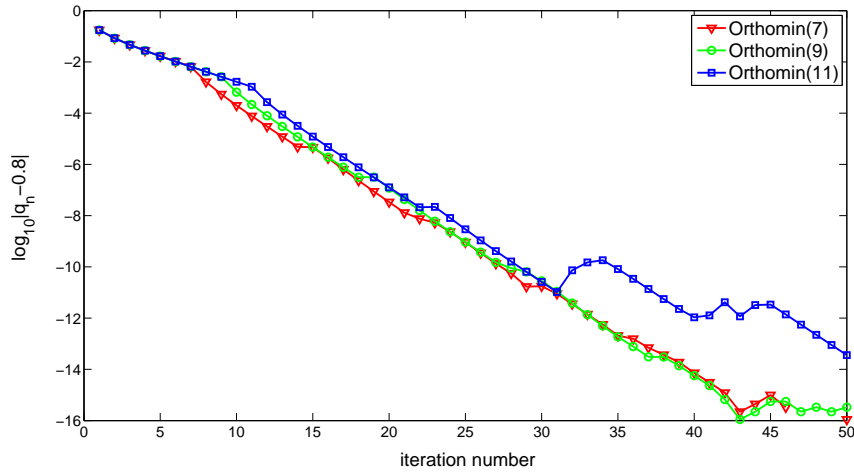


FIG. 2.2.  $\log_{10} |q_n - \rho|$  for *Orthomin*( $k$ ),  $k = 7, 9, 11$  ( $\rho = 0.8$ ,  $d = 32$ ).

be a uniform grid (we identify  $x_0$  with  $x_d$ ,  $x_{-1}$  with  $x_{d-1}$ , and  $x_1$  with  $x_{d+1}$ ), and

replace the derivatives in (2.10) with the usual centered difference formulas

$$-u''(x_j) \approx \frac{2u(x_j) - u(x_{j-1}) - u(x_{j+1}))}{h^2}, \quad u'(x_j) \approx \frac{u(x_{j+1}) - u(x_{j-1}))}{2h}.$$

The resulting discretization<sup>2</sup> is a linear system of type (1.1):  $A$  is a normal matrix with orthogonal eigenvectors  $\chi^{(k)} \in \mathbb{C}^d$  and corresponding eigenvalues  $\lambda_k$  given by

$$\chi_j^{(k)} = \exp(\mathbf{i}kjh), \quad \lambda_k = -\frac{2a}{h^2} \cos(kh) + \mathbf{i} \frac{b}{h} \sin(kh) + c + \frac{2a}{h^2}.$$

The eigenvalues lie on an ellipse with semi-axes  $2a/h^2$  and  $b/h$ ; when  $2a = hb$  this is a circle of radius  $b/h$ . After further rescaling, the system can be brought to the form (2.1). However, as will be shown in Section 3, this example is very relevant to the convergence study of Orthomin( $k$ ) also when  $2a \neq hb$ .

**3. Further examples: normal matrices with spectra on ellipses.** So far we have examined the systems (2.1), and we conjectured that for any  $k \in \mathbb{N}$  we can find operators  $U$  of the form (2.6) so that for all  $1 \leq j \leq k$ , Orthomin( $j$ ) achieves an asymptotic convergence rate equal to  $\rho$ . After a trivial rescaling, we restate Conjecture 2.2 in the following way: for any circle  $\mathcal{C}$  of center  $z_0$  and radius  $\rho$  satisfying  $0 < \rho < |z_0|$  there exists a normal matrix  $A$  whose spectrum lies on  $\mathcal{C}$  so that for all  $1 \leq j \leq k$ , the Orthomin( $j$ ) iteration applied to the system (1.1) with  $b = (1, 1, \dots, 1)^T$  and zero initial guess has an asymptotic convergence rate of  $\rho/|z_0|$ .

In this section we show numerical evidence suggesting that if we replace the circle  $\mathcal{C}$  with a non-circular ellipse  $\mathcal{E}$  in the example above, all Orthomin( $j$ ) with  $k \geq 2$  achieve **the same** asymptotic convergence rate  $\rho_{\mathcal{E}}$ , which is smaller than the asymptotic convergence rate of Orthomin(1). For the exact formulation see Conjecture 3.1. We remark that the discretized numerical PDE from Section 2.4 is an example of precisely such a system.

In order to make the examples very specific we first describe an ellipse  $\mathcal{E}$  by its semi-axes  $\alpha > 0$  and  $\beta > 0$ , the angle  $\theta \in \mathbb{R}$  between its axes and the coordinate axes, and the position  $u \in \mathbb{C}$  of its center:

$$\mathcal{E} = \{u + e^{\mathbf{i}\theta}(\alpha \cos \gamma + \mathbf{i}\beta \sin \gamma) : \gamma \in [0, 2\pi]\} . \quad (3.1)$$

It is assumed that neither  $\mathcal{E}$  nor its interior contain the origin. For  $d \in \mathbb{N}$  we consider the numbers  $\mu_j \in \mathcal{E}$  defined as

$$\mu_j = u + e^{\mathbf{i}\theta} \left( \alpha \cos \frac{2\pi j}{d} + \mathbf{i}\beta \sin \frac{2\pi j}{d} \right), \quad j = 1, \dots, d. \quad (3.2)$$

As before, we associate a linear operator

$$A_{\mathcal{E},d} \stackrel{\text{def}}{=} \text{diag}[\mu_1, \dots, \mu_d].$$

To construct an analogous continuous-space operator let  $d\mu_{\mathcal{E}}(z)$  denote the arc-length measure on  $\mathcal{E}$ . Define

$$A_{\mathcal{E}} : L^2(\mathcal{E}, d\mu_{\mathcal{E}}) \rightarrow L^2(\mathcal{E}, d\mu_{\mathcal{E}}), \quad A_{\mathcal{E}} f(z) = z f(z). \quad (3.3)$$

**CONJECTURE 3.1.** *For any ellipse  $\mathcal{E}$  there exists a number  $\rho_{\mathcal{E}} \in (0, 1)$  so that the following hold:*

---

<sup>2</sup>This particular discretization is not appropriate for advection-dominated problems.

- (i) For all  $k \in \mathbb{N}$  with  $k \geq 2$ , there exists  $d_k \in \mathbb{N}$  so that for  $d \geq d_k$  the ratio  $q_n = \|r_{n+1}^{(k)}\|/\|r_n^{(k)}\|$  of the residual-norm obtained by applying the Orthomin( $k$ ) iteration with zero initial guess to the system

$$A_{\mathcal{E},d}x = (1, 1, \dots, 1)^T \quad (3.4)$$

satisfies

$$\lim_{n \rightarrow \infty} q_n = \rho_{\mathcal{E}} . \quad (3.5)$$

- (ii) For any  $k \geq 2$  the same ratio obtained by applying the Orthomin( $k$ ) iteration with zero initial guess to the continuous system

$$A_{\mathcal{E}}x = 1$$

also satisfies (3.5).

- (iii) If the ellipse is not circular, then  $\rho_{\mathcal{E}}$  is smaller than the asymptotic convergence rate of Orthomin(1).

Two facts are notable about the behavior of Orthomin( $k$ ) on the systems in Conjecture 3.1. First, it is remarkable that the ratios  $q_n$  converge at all; indeed, we show that for  $k = 1$  the sequence  $\{q_n\}_{n \in \mathbb{N}}$  is convergent regardless of the choice of the numbers  $\mu_1, \dots, \mu_d$ , but for  $k \geq 2$  the sequence  $\{q_n\}_{n \in \mathbb{N}}$  may not be monotone, and is not expected to converge in general. The second interesting fact is that all Orthomin( $k$ ) with  $k \geq 2$  achieve the same asymptotic convergence rate for sufficiently large  $d$ . Moreover, numerical experiments show that  $q_n$  converges to the same limit  $\rho_{\mathcal{E}}$  even for a random initial guess and right-hand side  $b$ . However, in spite of the fact that  $\rho_{\mathcal{E}}$  seems to be intimately related to the ellipse, currently we do not understand the nature of this connection, i.e., how to compute  $\rho_{\mathcal{E}}$  using only information about  $\mathcal{E}$ .

We conclude this section by showing numerical evidence in support of Conjecture 3.1. For numerical experiments we have selected an ellipse in general position (not aligned with the coordinate axes) with  $\alpha = 2$ ,  $\beta = 1$ ,  $u = 2 + \mathbf{i}$ , and  $\theta = \pi/6$  (see Figure 3.1). For  $d = 128$  we solved the system (3.4) using Orthomin( $k$ ) with  $k = 1, 2, 3, 4, 10$ . In Figure 3.2 we plot the ratios  $q_n$  for each of the solves. The data strongly suggests that for  $k = 2, 3, 4, 10$  we have

$$q_n \rightarrow \rho_{\mathcal{E}} \approx 0.6891227 .$$

This approximate value (up to the first eight digits) was also obtained when solving (3.4) with random right-hand side and initial guess. In the particular case of Orthomin(1), we know that  $q_n$  is convergent (and monotone increasing): numerically we find that  $\lim q_n \approx 0.7902$ .

**4. Convergence analysis for Orthomin(1).** The main objective of this section is to prove Conjectures 2.2 and 2.3 for  $k = 1$ . In Section 4.1 we show that the sequence  $\{q_n\}_{n \in \mathbb{N}}$  is increasing and bounded. After stating in Section 4.2 a few technical results, we discuss in Section 4.3 examples when  $q_n$  does not converge to  $\rho$ . The behavior of  $q_n$  for two-dimensional systems is presented in Section 4.4. In Section 4.5 we prove Conjecture 2.2 for  $k = 1$ . Sections 4.6 and 4.7 are devoted to the infinite-dimensional case (Conjecture 2.3).

We consider matrices of the form

$$A = \text{diag}[\mu_1, \dots, \mu_d] , \quad (4.1)$$



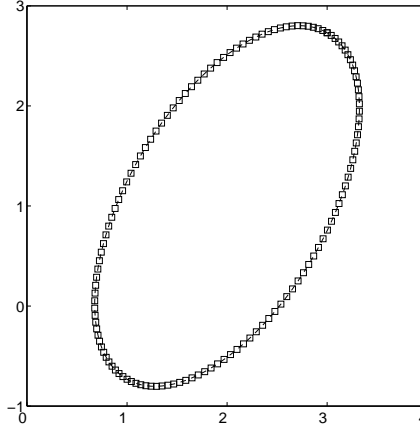


FIG. 3.1. The ellipse with semi-axes  $\alpha = 2$ ,  $\beta = 1$ ,  $u = 2 + i$ , and  $\theta = \pi/6$  ( $d = 128$ ).

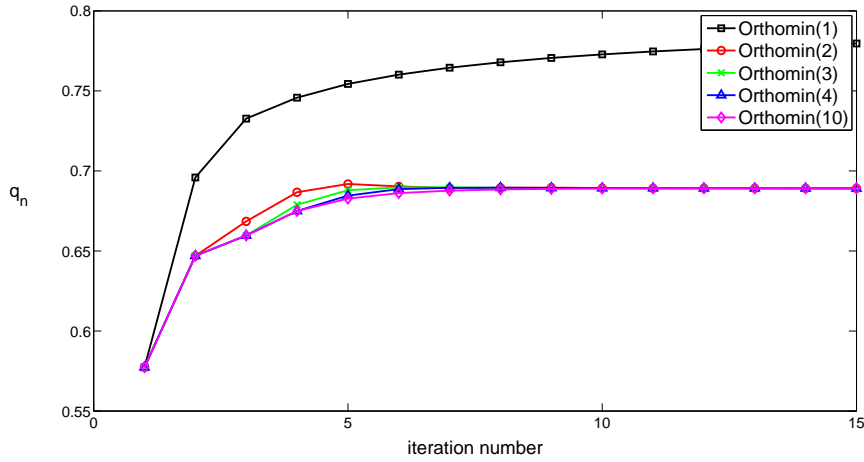


FIG. 3.2. The comparative residual norms for Orthomin( $k$ ) ( $k = 1, 2, 3, 4, 5, 10$ ): for Orthomin(1)  $q_n$  exceeds 0.7902, but for  $k = 2, 3, 4, 5, 10$  we note a convergence of  $q_n$  to a value near 0.6891227.

with  $\mu_1, \dots, \mu_d \in \mathbb{C}$  nonzero complex numbers. Since we are interested in the evolution of the residuals, we retain only the recursive equation from Orthomin(1) that produces the residual  $r_n = r_n^{(1)}$ :

$$r_{n+1} = r_n - \Pi_{Ar_n} r_n, \quad (4.2)$$

with  $r_0 \in \mathbb{C}^d$  being chosen arbitrarily. Recall from (1.4) and (1.6) that

$$\lambda_n = \frac{(r_n, Ar_n)}{(Ar_n, Ar_n)}, \quad r_{n+1} = r_n - \lambda_n Ar_n.$$

Let  $r_n = (r_n^1, \dots, r_n^d)$  be the coefficients of  $r_n$ . We consider the finite probability measure supported at  $1, \dots, d$  with weights proportional to  $|r_n^1|^2, \dots, |r_n^d|^2$ . We will

refer to it as the  $r_n$ -measure, and use the subscript  $n$  to denote it. For instance, the expected value of a vector  $\xi = (\xi_1, \dots, \xi_d)$  with respect to this measure is

$$\mathbb{E}_n(\xi) := \frac{\sum_{k=1}^d \xi_k |r_n^k|^2}{\sum_{k=1}^d |r_n^k|^2}.$$

Since  $r_{n+1} = r_n - \lambda_n A r_n$  has coefficients  $r_{n+1}^k = (1 - \lambda_n \mu_k) r_n^k$ , the following change of variable formula holds:

$$\mathbb{E}_{n+1}(\xi) = \frac{\mathbb{E}_n(\xi |1 - \lambda_n \mu|^2)}{\mathbb{E}_n(|1 - \lambda_n \mu|^2)}, \quad (4.3)$$

where  $\mu = (\mu_1, \dots, \mu_d)$  is the vector of eigenvalues of  $A$ . In particular,

$$\lambda_n = \frac{\sum \bar{\mu}_k |r_n^k|^2}{\sum |\mu_k|^2 |r_n^k|^2} = \frac{\mathbb{E}_n(\bar{\mu})}{\mathbb{E}_n(|\mu|^2)}. \quad (4.4)$$

**4.1. Monotonocity of  $q_n$ .** We first show that the sequence  $q_n$  is increasing and bounded.

**PROPOSITION 4.1.** *If  $r_n$  is given by (4.2) and  $A$  is defined as in (4.1), then  $q_n$  is increasing and bounded between 0 and 1.*

*Proof.* We will use the measure-theoretic notation:

$$\begin{aligned} q_n^2 &= \frac{\|r_{n+1}\|^2}{\|r_n\|^2} = \mathbb{E}_n(|1 - \lambda_n \mu|^2) = \mathbb{E}_n(1 + |\lambda_n|^2 |\mu|^2 - \lambda_n \mu - \bar{\lambda}_n \bar{\mu}) \\ &= 1 + |\lambda_n|^2 \mathbb{E}_n(|\mu|^2) - \lambda_n \mathbb{E}_n(\mu) - \bar{\lambda}_n \mathbb{E}_n(\bar{\mu}) \stackrel{(4.4)}{=} 1 - \frac{|\mathbb{E}_n(\mu)|^2}{\mathbb{E}_n(|\mu|^2)}. \end{aligned}$$

We compare  $1 - q_{n+1}^2$  and  $1 - q_n^2$ . For the latter, we use the change of variable formula (4.3):

$$1 - q_{n+1}^2 = \frac{|\mathbb{E}_{n+1}(\mu)|^2}{\mathbb{E}_{n+1}(|\mu|^2)} = \frac{|\mathbb{E}_n(\mu |1 - \lambda_n \mu|^2)|^2}{\mathbb{E}_n(|1 - \lambda_n \mu|^2) \mathbb{E}_n(|\mu|^2 |1 - \lambda_n \mu|^2)}.$$

We can re-write this as

$$1 - q_{n+1}^2 = \frac{|\mathbb{E}_n(\xi |1 - \xi|^2)|^2}{\mathbb{E}_n(|1 - \xi|^2) \mathbb{E}_n(|\xi|^2 |1 - \xi|^2)},$$

with  $\xi = \lambda_n \mu$ . By construction,

$$\mathbb{E}_n(\xi) = \mathbb{E}_n(|\xi|^2) = \frac{|\mathbb{E}_n(\mu)|^2}{\mathbb{E}_n(|\mu|^2)},$$

hence we can apply the result of Lemma 4.2 to  $\xi$ :

$$1 - q_n^2 = \mathbb{E}_n(|\xi|^2) \geq \frac{|\mathbb{E}_n(\xi |1 - \xi|^2)|^2}{\mathbb{E}_n(|1 - \xi|^2) \mathbb{E}_n(|\xi|^2 |1 - \xi|^2)} = 1 - q_{n+1}^2,$$

hence  $q_n \leq q_{n+1}$ .  $\square$

**LEMMA 4.2.** *Let  $\xi$  a complex-valued random variable with finite moments up to order 4 satisfying the identity  $\mathbb{E}(\xi) = \mathbb{E}(|\xi|^2)$ . The following inequality then holds:*

$$\mathbb{E}(|\xi|^2) \mathbb{E}(|1 - \xi|^2) \mathbb{E}(|\xi|^2 |1 - \xi|^2) \geq |\mathbb{E}(\xi |1 - \xi|^2)|^2. \quad (4.5)$$

*Proof.* First of all, we remark that if  $\xi$  satisfies the condition stated in the Lemma, then so does  $1 - \xi$ . Thus, the situation is symmetric in  $\xi$  and  $1 - \xi$ .

Let  $\theta \in \mathbb{R}$  such that  $\mathbb{E}(\xi|1 - \xi|^2) = e^{i\theta}|\mathbb{E}(\xi|1 - \xi|^2)|$ . Consider the function

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(t) := \text{Var}(t(1 - \xi) + e^{i\theta}\bar{\xi}(1 - \xi)), \quad (4.6)$$

where  $\text{Var}(\xi) = \mathbb{E}(|\xi|^2) - |\mathbb{E}(\xi)|^2$  denotes the variance of a random variable  $\xi$ . By opening up the parenthesis inside the expected value, we obtain

$$\begin{aligned} f(t) &= t^2\{\mathbb{E}(|1 - \xi|^2) - |\mathbb{E}(1 - \xi)|^2\} + 2t|\mathbb{E}(\xi|1 - \xi|^2)| + \mathbb{E}(|\xi|^2|1 - \xi|^2) \\ &= t^2\mathbb{E}(|\xi|^2)\mathbb{E}(|1 - \xi|^2) + 2t|\mathbb{E}(\xi|1 - \xi|^2)| + \mathbb{E}(|\xi|^2|1 - \xi|^2). \end{aligned}$$

The second equality follows from a manipulation of the coefficient of  $t^2$  which takes into account the fact that  $\mathbb{E}(\xi) = \mathbb{E}(|\xi|^2)$ . This shows that  $f(t)$  is a real valued quadratic form. The fact that it is a positive definite quadratic form follows from the fact that the variance of a random variable is always a positive number. Therefore,  $f(t)$  has negative discriminant:

$$|\mathbb{E}(\xi|1 - \xi|^2)|^2 - \mathbb{E}(|\xi|^2)\mathbb{E}(|1 - \xi|^2)\mathbb{E}(|\xi|^2|1 - \xi|^2) \leq 0. \quad \square$$

We should point out that in the case when  $\xi$  is real valued (which is not the case we are dealing with), the statement of Lemma 4.2 can be reduced to Pearson's inequality [9] (see also [6]) between the skewness  $\tau$  and the kurtosis  $\kappa$  of a distribution:  $\kappa - \tau^2 - 1 \geq 0$ . We do not give a proof of this fact, as it is of no relevance to the rest of the paper. Note that we can think of  $q_n$  as measuring the dispersion of the random variable  $\mu$  relative to the  $r_n$  measure: variance about the mean divided by average size. The monotonicity of  $q_n$  reflects the fact that  $\mu$  becomes increasingly more uniformly distributed relative to the  $r_n$ -measures.

It is important to remark that Proposition 4.1 holds for all normal (non-singular) matrices since they can be diagonalized using unitary transformations which leave the residual norms of  $\text{Orthomin}(k)$ , and hence  $q_n$ , unchanged.

**4.2. The case  $\mu_k = 1 + \rho\zeta_k$ ,  $|\zeta_k| = 1$ , and  $r_0 \in \mathbb{C}^d$  arbitrary.** In this section we assume that  $A$  is of the form  $A = I + \rho U$ ,  $U = \text{diag}[\zeta_1, \dots, \zeta_d]$ , with  $0 < \rho < 1$  and  $|\zeta_1| = \dots = |\zeta_d| = 1$ . Also, we keep  $r_0 \in \mathbb{C}^d$  arbitrary unless otherwise specified. We introduce the following quantities, for  $n \geq 0$ :

$$\omega_n = \frac{(Ur_n, r_n)}{(r_n, r_n)}, \quad T_n = \frac{1 - \lambda_n}{\rho\lambda_n}. \quad (4.7)$$

Note that the coefficients of  $r_{n+1}$  are related to those of  $r_n$  as follows

$$r_{n+1}^k = (1 - \lambda_n\mu_k)r_n^k = \rho\lambda_n(T_n - \zeta_k)r_n^k, \quad (4.8)$$

and the change of variable formula becomes

$$\mathbb{E}_{n+1}(\xi) = \frac{\mathbb{E}_n(\xi|T_n - \zeta|^2)}{\mathbb{E}_n(|T_n - \zeta|^2)}. \quad (4.9)$$

LEMMA 4.3. *For  $n \geq 0$  we have*

$$\lambda_n = \frac{1 + \rho\bar{\omega}_n}{1 + \rho^2 + 2\rho R\omega_n}, \quad T_n = \frac{\omega_n + \rho}{\rho\bar{\omega}_n + 1}, \quad q_n^2 = \rho^2 \frac{1 - |\omega_n|^2}{1 + \rho^2 + 2\rho R\omega_n}, \quad (4.10)$$

where  $\Re z$  denotes the real part of a complex number  $z$ .

*Proof.* Let  $\zeta = (\zeta_1, \dots, \zeta_d)$ . Clearly,  $\omega_n = \mathbb{E}_n(\zeta)$ . Since  $\mu = 1 + \rho\zeta$ , we have

$$\lambda_n = \frac{\mathbb{E}_n(1 + \rho\bar{\zeta})}{\mathbb{E}_n(|1 + \rho\zeta|^2)}.$$

The formula for  $\lambda_n$  then follows from the fact that  $\mathbb{E}_n(1 + \rho\bar{\zeta}) = 1 + \rho\bar{\omega}_n$ , and  $\mathbb{E}_n(|1 + \rho\zeta|^2) = 1 + \rho^2 + 2\rho\Re\omega_n$ . Next, the formula of  $T_n$  is a direct consequence of the formula of  $\lambda_n$ . Finally,

$$q_n^2 = 1 - \frac{|\mathbb{E}_n(\mu)|^2}{\mathbb{E}_n(|\mu|^2)} = 1 - \frac{1 + \rho^2|\omega_n|^2 + 2\rho\Re\omega_n}{1 + \rho^2 + 2\rho\Re\omega_n} = \frac{\rho^2(1 - |\omega_n|^2)}{1 + \rho^2 + 2\rho\Re\omega_n}. \quad \square$$

PROPOSITION 4.4. *For  $n \geq 0$  we have  $|\omega_n| \leq 1$  and  $0 \leq q_n \leq \rho$ . Moreover, the following statements are equivalent:*

$$(a) \lim_{n \rightarrow \infty} q_n = \rho. \quad (b) \lim_{n \rightarrow \infty} \omega_n = -\rho. \quad (c) \lim_{n \rightarrow \infty} \lambda_n = 1. \quad (d) \lim_{n \rightarrow \infty} T_n = 0. \quad (4.11)$$

*Proof.* The bound  $|\omega_n| \leq 1$  follows from  $\|U\| \leq 1$ . The fact that  $q_n$  is increasing has been proved in the previous section, and the bound  $q_n \leq \rho$  is a direct consequence of Theorem 2.1. Since  $\lambda_n, T_n, q_n$  are continuous functions of  $\omega_n$ , the statement (b) clearly implies all the others. We also have (a)  $\Rightarrow$  (b) since  $\frac{1 - |\omega|^2}{1 + \rho^2 + 2\rho\Re\omega} \leq 1$ , with equality for  $\omega = -\rho$ . Similarly (d)  $\Rightarrow$  (b) since  $T_n$  has bounded denominator. Finally

$$1 - \lambda_n = \frac{\rho(\rho + \omega_n)}{1 + \rho^2 + 2\rho\Re\omega_n}.$$

Since the denominator is bounded,  $\lim_n \lambda_n = 1$  implies  $\lim_n \omega_n = -\rho$  ((c)  $\Rightarrow$  (b)).  $\square$

In addition to the quantities defined above, we also define the higher moments

$$\omega_{n,j} := \frac{(U^j r_n, r_n)}{(r_n, r_n)} = \mathbb{E}_n(\zeta^j), \quad j \geq 0. \quad (4.12)$$

These moments are used in the convergence proof in Section 4.5. Clearly,  $\omega_{n,0} = 1$  and  $\omega_{n,1} = \omega_n$ . Using the change of variable formula (4.9),

$$\omega_{n+1,j} = \mathbb{E}_{n+1}(\zeta^j) = \frac{\mathbb{E}_n(\zeta^j |T_n - \zeta|^2)}{\mathbb{E}_n(|T_n - \zeta|^2)} = \frac{\mathbb{E}_n\{\zeta^j(1 + |T_n|^2 - T_n\bar{\zeta} - \bar{T}_n\zeta)\}}{\mathbb{E}_n\{1 + |T_n|^2 - T_n\bar{\zeta} - \bar{T}_n\zeta\}},$$

and we obtain the following

PROPOSITION 4.5. *For  $n \geq 0$  we have the following recurrence relation*

$$\omega_{n+1,j} = \frac{(1 + |T_n|^2)\omega_{n,j} - T_n\omega_{n,j-1} - \bar{T}_n\omega_{n,j+1}}{1 + |T_n|^2 - 2\Re(\bar{T}_n\omega_n)}, \quad j \geq 1. \quad (4.13)$$

COROLLARY 4.6. *The collection of moments of the initial distribution  $\{\omega_{0,j}\}_{j \geq 0}$  completely determine the sequence  $\omega_n$  and implicitly  $q_n$ .*

Note that when  $\zeta_k$  are the roots of unity of order  $d$  we have  $\omega_{n,j+d} = \omega_{n,j}$ ; hence the finite set of initial moments  $\{\omega_{0,j}\}_{1 \leq j \leq d}$  completely determine the sequence  $q_n$ .

**4.3. Non-convergence to  $\rho$ .** Let  $\text{Hull}(\zeta_1, \dots, \zeta_d)$  denote the convex hull of  $\zeta_1, \dots, \zeta_d$ . This is a compact convex subset of  $\mathbb{C}$ . Since

$$\omega_n = \frac{\sum_{k=1}^d \zeta_k |r_n(k)|^2}{\sum_{k=1}^d |r_n(k)|^2} \in \text{Hull}(\zeta_1, \dots, \zeta_d),$$

the sequence  $\omega_n$  cannot converge to  $-\rho$  unless  $-\rho \in \text{Hull}(\zeta_1, \dots, \zeta_d)$ . Since the statements  $\lim_n \omega_n = -\rho$  and  $\lim_n q_n = \rho$  are equivalent, we have the following.

**PROPOSITION 4.7.** *Assume  $-\rho \notin \text{Hull}(\zeta_1, \dots, \zeta_d)$ . Then  $\lim_n q_n \neq \rho$ .*

**COROLLARY 4.8.** *Assume that  $\rho \in (0, 1)$  is arbitrary, and  $|\theta_k| < \pi - \arccos(\rho)$ , for  $1 \leq k \leq d$ . If  $\zeta_k = \exp(i\theta_k)$ , then  $\lim q_n \neq \rho$ .*

*Proof.* The angles are chosen so that  $\Re(\zeta_k) > \rho$ . This ensures  $-\rho \notin \text{Hull}(\zeta_1, \dots, \zeta_d)$ , and the previous Proposition applies.  $\square$

Figures 4.1 and 4.2 illustrate the context of Corollary 4.8:  $q_n$  does not converge to  $\rho$ , and  $\omega_n$  does not converge to  $-\rho$ . We end this section with a sharpened version of Conjecture 2.2 for  $k = 1$ :

**CONJECTURE 4.9.** *For  $\text{Orthomin}(1)$ , if  $-\rho \in \text{Hull}(\zeta_1, \dots, \zeta_d)$ , then  $q_n \rightarrow \rho$ .*

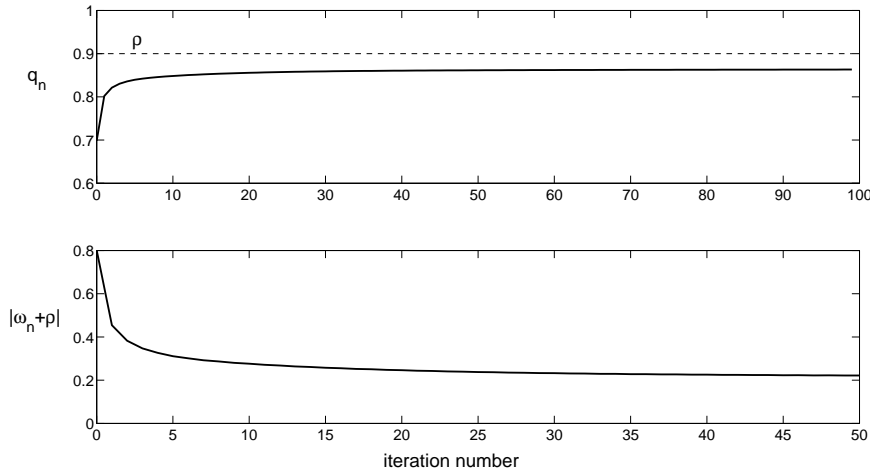


FIG. 4.1. The case when  $-\rho$  does not belong to  $\text{Hull}(\zeta_1, \dots, \zeta_d)$ :  $\rho = 0.9$ ,  $d = 15$ .

**4.4. The case  $d=2$ .** Surprisingly, this case is not completely trivial either.

**PROPOSITION 4.10.** *Assume  $d = 2$  and the initial vector  $r_0 \in \mathbb{C}^2$  is arbitrary, with non-zero entries. Then  $q_n$  is a constant depending on  $r_0$ , while  $\omega_n$  is a periodic sequence with period 2. The convergence (2.7) for  $k = 1$  does not hold in this case.*

*Proof.* With a rotation, we may assume  $\zeta_1 = 1$  and  $\zeta_2 = \zeta$  is arbitrary. Then  $\mu_1 = 1 + \rho$  and  $\mu_2 = 1 + \rho\zeta$ . We have

$$\lambda_0 = \frac{(r_0, Ar_0)}{(Ar_0, Ar_0)} = \frac{\bar{\mu}_1 |r_0^1|^2 + \bar{\mu}_2 |r_0^2|^2}{\|Ar_0\|^2}, \quad \omega_0 = \frac{|r_0^1|^2 + \zeta |r_0^2|^2}{|r_0^1|^2 + |r_0^2|^2},$$

therefore

$$1 - \lambda_0 \mu_1 = \frac{-\rho(1 - \zeta) \bar{\mu}_2 |r_0^2|^2}{\|Ar_0\|^2}, \quad 1 - \lambda_0 \mu_2 = \frac{\rho(1 - \zeta) \bar{\mu}_1 |r_0^1|^2}{\|Ar_0\|^2}.$$

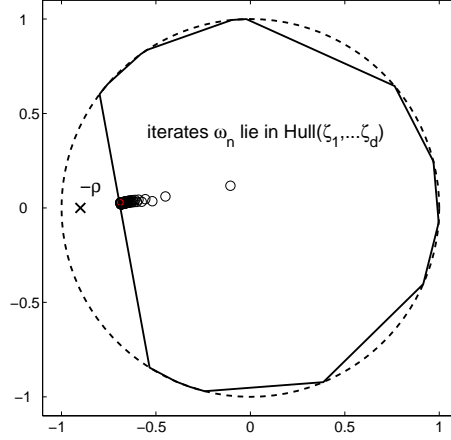


FIG. 4.2. The case when  $-\rho$  does not belong to  $\text{Hull}(\zeta_1, \dots, \zeta_d)$ :  $\rho = 0.9$ ,  $d = 15$ .

On the other hand  $r_1 = r_0 - \lambda_0 A r_0$ , hence  $r_1^1 = (1 - \lambda_0 \mu_1) r_0^1$ , and  $r_1^2 = (1 - \lambda_0 \mu_2) r_0^2$ . Therefore

$$\frac{r_1^2}{r_1^1} = \frac{1 - \lambda_0 \mu_2}{1 - \lambda_0 \mu_1} \cdot \frac{r_0^2}{r_0^1} = \frac{-\bar{\mu}_1 |r_0^1|^2}{\bar{\mu}_2 |r_0^2|^2} \frac{r_0^2}{r_0^1} = \frac{-\bar{\mu}_1}{\bar{\mu}_2} \frac{\bar{r}_0^1}{\bar{r}_0^2} \Rightarrow \frac{|r_1^2|}{|r_1^1|} = \frac{|\mu_1|}{|\mu_2|} \frac{|r_0^1|}{|r_0^2|}. \quad (4.14)$$

By applying the same procedure to  $r_1$  instead of  $r_0$ , we obtain

$$\frac{|r_2^2|}{|r_2^1|} = \frac{|r_0^2|}{|r_0^1|} \Rightarrow \frac{|r_2^1|^2 + \zeta |r_2^2|^2}{|r_2^1|^2 + |r_2^2|^2} = \frac{|r_0^1|^2 + \zeta |r_0^2|^2}{|r_0^1|^2 + |r_0^2|^2}, \quad \text{i.e. } \omega_2 = \omega_0.$$

This shows that the sequence  $\omega_n$  is periodic with period 2. With the above formulae for  $1 - \lambda \mu_1$  and  $1 - \lambda \mu_2$ , we also have

$$\frac{\|r_1\|^2}{\|r_0\|^2} = \frac{|1 - \lambda \mu_1|^2 |r_0^1|^2 + |1 - \lambda \mu_2|^2 |r_0^2|^2}{|r_0^1|^2 + |r_0^2|^2} = \rho^2 |1 - \zeta|^2 \frac{|r_0^1|^2 |r_0^2|^2}{\|r_0\|^2 \|A r_0\|^2}. \quad (4.15)$$

Let  $y = |r_0^2|^2 / |r_0^1|^2$ . The above fraction equals, up to a constant,

$$\frac{1}{1 + 1/y} \cdot \frac{1}{1 + y |\mu_2|^2 / |\mu_1|^2} =: g(y).$$

Because of (4.14), substituting  $r_1$  for  $r_0$  amounts to substituting  $y$  by  $\frac{|\mu_1|^2}{|\mu_2|^2} \frac{1}{y}$ . This does not change the value of  $g(y)$ , which means that  $\frac{\|r_2\|^2}{\|r_1\|^2} = \frac{\|r_1\|^2}{\|r_0\|^2}$ . This proves that  $q_2 = q_1$ . Similarly,  $q_n = q_{n-1}$  for  $n \geq 2$ .  $\square$

**4.5. Convergence of  $q_n$  to  $\rho$ .** We have already seen that  $\lim_n q_n = \rho$  if and only if  $\lim_n \lambda_n = 1$ . In this section we will work with the quantities

$$\beta_n := 1 - \lambda_n, \quad u_n := \omega_{n,1}, \quad \text{and } v_n := \omega_{n,2},$$

and we formulate sufficient conditions that guarantee  $\beta_n \rightarrow 0$ . We have

$$r_{n+1} = r_n - \lambda_n (I + \rho U) r_n = \beta_n A r_n - \rho U r_n, \quad (4.16)$$

and

$$(r_{n+1}, r_{n+1}) = (r_{n+1}, r_n - \lambda_n Ar_n) = (r_{n+1}, r_n) . \quad (4.17)$$

Further, since  $U$  is unitary,

$$(1 - \rho) \leq \|A\| \leq (1 + \rho) . \quad (4.18)$$

Now,

$$\begin{aligned} 1 - \lambda_{n+1} &= 1 - \frac{(r_{n+1}, Ar_{n+1})}{\|Ar_{n+1}\|^2} = \rho \frac{(Ur_{n+1}, Ar_{n+1})}{\|Ar_{n+1}\|^2} \\ &= \rho \frac{(Ur_{n+1}, r_{n+1}) + \rho(r_{n+1}, r_{n+1})}{\|Ar_{n+1}\|^2} = \rho \frac{(Ur_{n+1}, r_{n+1}) + \rho(r_{n+1}, r_n)}{\|Ar_{n+1}\|^2} \\ &= \rho \frac{(Ur_{n+1}, r_{n+1}) + (Ur_{n+1}, \rho Ur_n)}{\|Ar_{n+1}\|^2} = \rho \frac{(Ur_{n+1}, r_{n+1} + \rho Ur_n)}{\|Ar_{n+1}\|^2} \\ &\stackrel{(4.16)}{=} \rho \frac{(Ur_{n+1}, (1 - \lambda_n) Ar_n)}{\|Ar_{n+1}\|^2} = \rho(1 - \bar{\lambda}_n) \frac{(Ur_{n+1}, Ar_n)}{\|Ar_{n+1}\|^2} . \end{aligned}$$

Therefore

$$\begin{aligned} \beta_{n+1} &= \rho \bar{\beta}_n \frac{(Ur_{n+1}, Ar_n)}{\|Ar_{n+1}\|^2} = \rho \bar{\beta}_n \frac{(U(\beta_n(I + \rho U)r_n - \rho Ur_n), (I + \rho U)r_n)}{\|Ar_{n+1}\|^2} \\ &= \rho \bar{\beta}_n \frac{\beta_n((U + \rho U^2)r_n, (I + \rho U)r_n) - (\rho U^2 r_n, (I + \rho U)r_n)}{\|Ar_{n+1}\|^2} \\ &= \rho \bar{\beta}_n (\beta_n((1 + \rho^2)u_n + \rho(1 + v_n)) - \rho^2 u_n - \rho v_n) \frac{\|r_n\|^2}{\|Ar_{n+1}\|^2} . \end{aligned}$$

Next, the statement

$$\|r_{n+1}\| = \|\beta_n Ar_n - \rho Ur_n\| \geq \rho \|Ur_n\| - \|\beta_n Ar_n\| \geq \|r_n\|(\rho - |\beta_n|(1 + \rho))$$

implies

$$\frac{\|r_n\|}{\|Ar_{n+1}\|} = \frac{\|r_n\|}{\|r_{n+1}\|} \frac{\|r_{n+1}\|}{\|Ar_{n+1}\|} \leq \frac{1}{(\rho - |\beta_n|(1 + \rho))(1 - \rho)} .$$

Therefore

$$|\beta_{n+1}| \leq \rho |\beta_n| \frac{(|\beta_n|((1 + \rho^2)|u_n| + \rho(1 + |v_n|)) + \rho^2|u_n| + \rho|v_n|)}{(\rho - |\beta_n|(1 + \rho))^2(1 - \rho)^2} . \quad (4.19)$$

Next we need to estimate  $|u_n|, |v_n|$ . We have

$$\begin{aligned} u_{n+1} &= \frac{(Ur_{n+1}, r_{n+1})}{(r_{n+1}, r_{n+1})} = \frac{(U(\beta_n Ar_n - \rho Ur_n), \beta_n Ar_n - \rho Ur_n)}{\|r_{n+1}\|^2} \\ &= \frac{|\beta_n|^2(AUr_n, Ar_n) - \rho(\beta_n(Ar_n, r_n) + \bar{\beta}_n(U^2 r_n, Ar_n)) + \rho^2(Ur_n, r_n)}{\|r_{n+1}\|^2} , \end{aligned}$$

hence

$$|u_{n+1}| = (|\beta_n|^2 \|A\|^2 + 2\rho |\beta_n| \cdot \|A\| + \rho^2 |u_n|) \frac{\|r_n\|^2}{\|r_{n+1}\|^2} \quad (4.20)$$

$$\leq \frac{|\beta_n|^2(1 + \rho)^2 + 2\rho |\beta_n|(1 + \rho) + \rho^2 |u_n|}{(\rho - |\beta_n|(1 + \rho))^2} . \quad (4.21)$$

The analogous inequality can be derived for  $v_n$ . We summarize the previous inequalities in

PROPOSITION 4.11. *The following recurrence relations hold:*

$$\begin{aligned} |\beta_{n+1}| &\leq \rho |\beta_n| \cdot \frac{|\beta_n|[(1+\rho^2)|u_n| + \rho(1+|v_n|)] + \rho^2|u_n| + \rho|v_n|}{[\rho - |\beta_n|(1+\rho)]^2(1-\rho)^2}, \\ |u_{n+1}| &\leq \frac{|\beta_n|^2(1+\rho)^2 + 2\rho|\beta_n|(1+\rho) + \rho^2|u_n|}{(\rho - |\beta_n|(1+\rho))^2}, \\ |v_{n+1}| &\leq \frac{|\beta_n|^2(1+\rho)^2 + 2\rho|\beta_n|(1+\rho) + \rho^2|v_n|}{(\rho - |\beta_n|(1+\rho))^2}. \end{aligned} \quad (4.22)$$

We will also need the following inequality which we state without proof.

LEMMA 4.12. *For  $|x| \leq 0.1$ ,  $\frac{1}{(1-x)^2} \leq 1 + Cx$ , with  $C = 2.5$ .*

PROPOSITION 4.13. *Assume the following:  $0 < \rho < 0.1$ , and  $\omega_{0,1} = \omega_{0,2} = \omega_{0,3} = 0$ . Then, for  $n \geq 1$ , we have:*

- (i)  $|u_n| \leq \rho + 2.7 \sum_{k=2}^n \rho^k \leq \rho + 3\rho^2$  ;
- (ii)  $|v_n| \leq 2.7 \sum_{k=2}^n \rho^k \leq 3\rho^2$  ;
- (iii)  $|\beta_n| \leq \rho^{n+2}$ .

*Proof.* We use the recurrence relations (4.13) to compute the first few terms in the sequences  $\beta_n, u_n, v_n$ .

$$\begin{aligned} T_0 &= \frac{\omega_{0,1} + \rho}{\rho\bar{\omega}_{0,1} + 1} = \rho, \quad \beta_0 = \frac{\rho T_0}{1 + \rho T_0} = \frac{\rho^2}{1 + \rho^2}, \\ u_1 &= \omega_{1,1} = \frac{(1 + |T_0|^2)\omega_{0,1} - T_0\bar{\omega}_{0,0} - \bar{T}_0\omega_{0,2}}{1 + |T_0|^2 - 2\Re(\bar{T}_0\omega_{0,1})} = \frac{-\rho}{1 + \rho^2}, \\ v_1 &= \omega_{1,2} = \frac{(1 + |T_0|^2)\omega_{0,2} - T_0\bar{\omega}_{0,1} - \bar{T}_0\omega_{0,3}}{1 + |T_0|^2 - 2\Re(\bar{T}_0\omega_{0,1})} = 0, \\ T_1 &= \frac{\omega_{1,1} + 1}{\rho\bar{\omega}_{1,1} + 1} = \rho^3, \quad \beta_1 = \frac{\rho T_1}{1 + \rho T_1} = \frac{\rho^4}{1 + \rho^4}. \end{aligned}$$

The inequalities in the proposition are thus true for  $n = 1$ , and we proceed by induction. We assume that the statements (i-iii) are true for some  $n \geq 1$ , and we prove that they hold for  $n + 1$  as well. For that, we rely on the inequalities of Proposition 4.11. We start with the inequality (iii):

$$\begin{aligned} |\beta_{n+1}| &\leq \rho^{n+3} \times \frac{\rho^{n+2}[(1+\rho^2)(\rho + 3\rho^2) + \rho(1+3\rho^2)] + \rho^2(\rho + 3\rho^2) + 3\rho^3}{[\rho - \rho^{n+2}(1+\rho)]^2(1-\rho)^2} \\ &= \rho^{n+3} \times \frac{\rho^{n+1}[(1+\rho^2)(1+3\rho) + 1+3\rho^2] + \rho(1+3\rho) + 3\rho}{[1 - \rho^{n+1}(1+\rho)]^2(1-\rho)^2} \\ &\leq \rho^{n+3} \times \frac{\rho^2[(1+\rho^2)(1+3\rho) + 1+3\rho^2] + \rho(1+3\rho) + 3\rho}{[1 - \rho^2(1+\rho)]^2(1-\rho)^2}. \end{aligned}$$

The fraction on the right hand side has numerator equal to  $4\rho + 5\rho^2 + 3\rho^3 + 4\rho^4 + 3\rho^5$ . This is easily seen to be less than 0.5, as  $0 < \rho < 0.1$ . On the other hand, the denominator is certainly greater than  $0.9^2 \times (1 - \frac{1.1}{100})^2 > 0.7$ . Therefore the fraction



on right hand side is less than 1, and  $|\beta_{n+1}| \leq \rho^{n+3}$ .

For inequality (ii),

$$\begin{aligned}
|u_{n+1}| &\leq \frac{\rho^{2(n+2)}(1+\rho)^2 + 2\rho^{n+3}(1+\rho) + \rho^2|u_n|}{[\rho - \rho^{n+2}(1+\rho)]^2} \\
&= \frac{\rho^{2(n+1)}(1+\rho)^2 + 2\rho^{n+1}(1+\rho) + |u_n|}{[1 - \rho^{n+1}(1+\rho)]^2} \\
&= \frac{x^2 + 2x + |u_n|}{(1-x)^2}, \quad \text{with } x = \rho^{n+1}(1+\rho), \\
&\leq (1+Cx)(x^2 + 2x + |u_n|), \quad \text{with } C = 2.5, \\
&= |u_n| + x[2 + C|u_n| + (2C+1)x + Cx^2].
\end{aligned}$$

From the induction step,  $|u_n| \leq \rho + 3\rho^2$ . Also,  $x = \rho^{n+1}(1+\rho) \leq \rho^2(1+\rho)$ . The quantity inside the square brackets is less than

$$2 + C(\rho + 3\rho^2) + (2C+1)\rho^2(1+\rho) + C\rho^4(1+\rho)^2.$$

As  $0 < \rho < 1$ , this is easily seen to be less than 2.5. Therefore,

$$|u_{n+1}| \leq |u_n| + 2.5(1+\rho)\rho^{n+1} < |u_n| + 2.7\rho^{n+1}.$$

Hence  $|u_{n+1}| \leq |u_1| + 2.7 \sum_{k=2}^{n+1} \rho^k$ . The exact same method is applied to  $v_{n+1}$ .  $\square$

**THEOREM 4.14.** *Assume the following hold:*

- (a)  $0 < \rho < 0.1$ ;
- (b)  $d \geq 4$ ;
- (c)  $r_0 = [1, \dots, 1]^T$ ;
- (d)  $\zeta_k$  are the roots of unity of order  $d$ ;
- (e)  $A = I + \rho U$ ,  $U = \text{diag}([\zeta_1, \dots, \zeta_d])$ .

*Then the sequence  $r_{n+1} = r_n - \Pi_{Ar_n} r_n$  satisfies*

$$\lim_{n \rightarrow \infty} \frac{\|r_{n+1}\|}{\|r_n\|} = \rho. \quad (4.23)$$

*Proof.* The hypotheses ensure that  $\omega_{0,1} = \omega_{0,2} = \omega_{0,3} = 0$ . Proposition 4.13 then applies to show

$$\lim_n \beta_n = 0 \Rightarrow \lim_n \lambda_n = 1 \stackrel{(4.11)}{\Rightarrow} \lim_n q_n = \rho. \quad \square$$

**4.6. The infinite dimensional case.** In this section we prove Conjecture 2.3 for  $k = 1$ . Recall that  $d\mu_0$  is the Haar probability measure on the unit circle  $\mathbb{S}^1$ , and assume that  $0 < \rho < 1$  is a fixed parameter (there will be no further restrictions on  $\rho$  in this section). For  $n \geq 0$ , we define

$$\omega_n := \frac{\int_{\mathbb{S}^1} z d\mu_n(z)}{\int_{\mathbb{S}^1} d\mu_n(z)}, \quad T_n := \frac{\omega_n + \rho}{\rho\bar{\omega}_n + 1}, \quad d\mu_{n+1}(z) := |z - T_n|^2 d\mu_n(z). \quad (4.24)$$

As shown in Section 4.2, the number  $\omega_n$  is equal to  $(Ur_n, r_n) / (r_n, r_n)$  where  $r_n$  is the residual obtained by applying Orthomin(1) to the system (1.1) with  $U$  chosen as in (2.8), and initial residual  $r_0 \equiv 1$ . Our goal is to show that  $\lim_n T_n = 0$  which is equivalent to Conjecture 2.3 for  $k = 1$ . For this, we need the following Lemma.

LEMMA 4.15. For  $q \in \mathbb{C}$  and  $n \geq 1$ , we define the polynomial in  $X$ :

$$\Phi_n(X, q) = \prod_{k=1}^n (X - q^{2k-1}) = \sum_{i=0}^n a_i X^i, \quad (4.25)$$

where  $a_i = a_i(q)$  are functions of  $q$ . The following identity holds:

$$\frac{\sum_{i=0}^{n-1} a_i a_{i+1}}{\sum_{i=0}^n a_i^2} = \frac{-q(1 - q^{2n})}{1 - q^{2n+2}}. \quad (4.26)$$

*Proof.* This identity can be proved using an identity of MacMahon (see [1, 7]) from the theory of partition functions. We give a few definitions to provide context. The  $q$ -Pochhammer symbol is defined as

$$(x; q)_n := \prod_{i=1}^n (1 - xq^{i-1}).$$

When  $|q| < 1$ , the product on the right hand side is convergent as  $n \rightarrow \infty$ , and the limit is denoted by  $(x; q)_\infty$ . The  $q$ -binomial coefficient is defined as

$$\begin{bmatrix} m \\ n \end{bmatrix}_q := \frac{(q; q)_n}{(q; q)_k (q; q)_{n-k}}.$$

MacMahon's identity ([7], paragraph 323) states that

$$(zq, q)_m (z^{-1}; q)_n = \sum_{k=-n}^m (-1)^k q^{k(k+1)/2} z^k \begin{bmatrix} m+n \\ n+k \end{bmatrix}_q. \quad (4.27)$$

Using our notation from (4.25), we note that

$$\Phi_n(x, q) \Phi_n(1/x, q) = (xq; q^2)_n (x^{-1}q; q^2)_n.$$

We apply (4.27) with  $q^2$  instead of  $q$  and  $z = xq^{-1}$ , to obtain

$$\Phi_n(x, q) \Phi_n(1/x, q) = \sum_{k=-n}^n (-1)^k q^{k^2} x^k \begin{bmatrix} 2n \\ n+k \end{bmatrix}_{q^2}. \quad (4.28)$$

Identifying the constant coefficient and the coefficient of  $x$  on both sides yields

$$\sum_{i=0}^n a_i^2 = \begin{bmatrix} 2n \\ n \end{bmatrix}_{q^2}, \quad \sum_{i=0}^{n-1} a_i a_{i+1} = -q \begin{bmatrix} 2n \\ n+1 \end{bmatrix}_{q^2}. \quad (4.29)$$

Therefore,

$$\begin{aligned} \frac{\sum_{i=0}^{n-1} a_i a_{i+1}}{\sum_{i=0}^n a_i^2} &= -q \begin{bmatrix} 2n \\ n+1 \end{bmatrix}_{q^2} / \begin{bmatrix} 2n \\ n \end{bmatrix}_{q^2} = -q \frac{(q^2; q^2)_{2n}}{(q^2; q^2)_{n-1} (q^2; q^2)_{n+1}} \frac{(q^2; q^2)_n^2}{(q^2; q^2)_{2n}} \\ &= \frac{-q(1 - q^{2n})}{1 - q^{2n+2}}. \quad \square \end{aligned}$$

THEOREM 4.16. Let  $\omega_n, T_n, \mu_n$  be defined by (4.24). Then  $T_n = \rho^{2n+1}$ ,  $\forall n \geq 0$ . In particular,  $\lim_n T_n = 0$ , which proves Conjecture 2.3 for  $k = 1$ .

*Proof.* We have

$$\omega_0 = \frac{\int z d\mu_0(z)}{\int d\mu_0(z)} = 0, \quad \text{hence } T_0 = \frac{0 + \rho}{0 \cdot d + 1} = \rho,$$

and the statement is true for  $n = 0$ . We proceed by induction. Let  $n \geq 1$  and assume that  $T_k = \rho^{2k+1}$  for  $k$  less than  $n - 1$ . From the definition,

$$d\mu_n(z) = \prod_{k=0}^{n-1} |z - \rho^{2k+1}|^2 d\mu_0(z) = |\Phi_n(z, \rho)|^2 d\mu_0(z). \quad (4.30)$$

With  $a_i = a_i(\rho)$ , the formula for  $\omega_n$  reads

$$\omega_n = \frac{\int z |\Phi_n(z, \rho)|^2 d\mu_0(z)}{\int |\Phi_n(z, \rho)|^2 d\mu_0(z)} = \frac{\int z |\sum_{i=0}^n a_i z^i|^2 d\mu_0}{\int |\sum_{i=0}^n a_i z^i|^2 d\mu_0}.$$

Since  $\int z^i d\mu_0(z)$  vanishes unless  $i = 0$ , we have

$$\omega_n = \frac{\sum_{i,j} a_i \bar{a}_j \int z^{i+1-j} d\mu_0(z)}{\sum_{i,j} a_i \bar{a}_j \int z^{i-j} d\mu_0(z)} = \frac{\sum_{i=0}^{n-1} a_i \bar{a}_{i+1}}{\sum_{i=0}^n |a_i|^2}. \quad (4.31)$$

We can now apply Lemma 4.15 (note that  $a_i(\rho)$  are real valued):

$$\omega_n = \frac{-\rho(1 - \rho^{2n})}{1 - \rho^{2n+2}}, \quad \text{and } T_n = \frac{\omega_n + \rho}{\rho \bar{\omega}_n + 1} = \rho^{2n+2},$$

which completes the induction step.  $\square$

*Remark.* The finite dimensional analogue of Theorem 4.16 is the case when  $d\mu_0$  is the equal weighted discrete probability measure supported at the roots of unity of order  $d$ . The exact same argument works there as well, but only up to  $n = d - 2$ . Beyond that, the right hand side of equation (4.31) is complicated by the higher correlation sums  $\sum_i a_i a_{i+kd}$ , with  $k \geq 1$ . This is due to the fact that, in the finite dimensional case, the integral  $\int z^n \mu_0(z) = \sum_{k=1}^d \exp(\frac{2\pi i n k}{d})$  can be nonzero beyond the trivial case  $n = 0$ : namely, when  $n \equiv 0 \pmod{d}$ . Our proof of Theorem 4.14 (which allows for some flexibility in the initial conditions) avoided the issue of explicitly computing  $T_n$ .

**4.7. Connection with the Jacobi Triple Product Formula.** Recall the Jacobi triple product formula ([1], Theorem 2.3) :

$$\prod_{k=1}^{\infty} (1 - z \rho^{2k-1})(1 - z^{-1} \rho^{2k-1})(1 - \rho^{2k}) = \sum_{k=-\infty}^{\infty} (-1)^k q^{k^2} z^k, \quad |\rho| < 1, |z| = 1. \quad (4.32)$$

When  $|z| = 1$  and  $\rho \in (0, 1)$ , this identity can be re-written as

$$\prod_{k=1}^{\infty} |z - \rho^{2k-1}|^2 = (\rho^2; \rho^2)_{\infty}^{-1} \sum_{k \in \mathbb{Z}} (-1)^k \rho^{k^2} z^k. \quad (4.33)$$

If we compare this to (4.30) we see that the sequence  $d\mu_n$  has a strong limit  $d\mu_{\infty}$ , whose density (with respect to the Haar measure  $d\mu_0$ ) is exactly (4.33).

Recall the quantities  $\omega_{n,k}$  defined at (4.12): in terms of the measures  $d\mu_n$ , they can be written as

$$\omega_{n,k} = \frac{\int_{\mathbb{S}^1} z^k d\mu_n(z)}{\int_{\mathbb{S}^1} d\mu_n(z)} . \quad (4.34)$$

We can see now that  $\lim_n \omega_{n,k}$  can be computed:

$$\lim_{n \rightarrow \infty} \omega_{n,k} = \frac{\int_{\mathbb{S}^1} z^k d\mu_\infty(z)}{\int_{\mathbb{S}^1} d\mu_\infty(z)} = (-1)^k \rho^{k^2} . \quad (4.35)$$

**Conclusions.** For  $k \in \mathbb{N}$  we give examples of linear systems (both finite and infinite-dimensional) for which we conjectured that  $\text{Orthomin}(1), \dots, \text{Orthomin}(k)$  achieve the same asymptotic convergence rate. These examples show that, in general,  $\text{Orthomin}(k)$  does not converge faster than  $\text{Orthomin}(1)$ . We analyze in detail the convergence of  $\text{Orthomin}(1)$  and provide numerical evidence in support of our conjectures with respect to  $\text{Orthomin}(k)$  for  $k > 1$ . The analysis for  $\text{Orthomin}(1)$  is fairly complicated and we do not see a straightforward way to extend the arguments to  $\text{Orthomin}(k)$  for  $k > 1$ . We provide numerical evidence that certain normal operators (related to numerical PDEs) with spectrum lying on an ellipse, have the following property:  $\text{Orthomin}(2)$ ,  $\text{Orthomin}(3)$ , etc. all have the same asymptotic convergence rate (depending only on the ellipse); moreover this is smaller than the asymptotic convergence rate of  $\text{Orthomin}(1)$ . This example offers a promising path to finding improved convergence rate estimates for  $\text{Orthomin}(2)$  under additional assumptions on the spectrum/field of values of the matrix. An important question, which remains unanswered, is whether there are applications where  $\text{Orthomin}(k)$ , perhaps coupled with preconditioners, can compete with the usual iterative solvers for non-symmetric systems.

## REFERENCES

- [1] GEORGE E. ANDREWS, *The theory of partitions*, Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1976. Encyclopedia of Mathematics and its Applications, Vol. 2.
- [2] STANLEY C. EISENSTAT, HOWARD C. ELMAN, AND MARTIN H. SCHULTZ, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (1983), pp. 345–357.
- [3] HOWARD C. ELMAN, *Iterative methods for large, sparse, nonsymmetric systems of linear equations*, PhD thesis, Dept. Computer Science, Yale Univ., New Haven, CT, 1982. Also available as Technical Report 229.
- [4] ANNE GREENBAUM, *Iterative methods for solving linear systems*, vol. 17 of Frontiers in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.
- [5] KARL E. GUSTAFSON AND DUGGIRALA K. M. RAO, *Numerical range*, Universitext, Springer-Verlag, New York, 1997.
- [6] CHRIS A. J. KLAASSEN, PHILIP J. MOKVELD, AND BERT VAN ES, *Squared skewness minus kurtosis bounded by 186/125 for unimodal distributions*, Statist. Probab. Lett., 50 (2000), pp. 131–135.
- [7] PERCY A. MACMAHON, *Combinatory analysis. Vol. I, II (bound in one volume)*, Dover Phoenix Editions, Dover Publications Inc., Mineola, NY, 2004. Reprint of *An introduction to combinatory analysis* (1920) and *Combinatory analysis. Vol. I, II* (1915, 1916).
- [8] P.K.W. VINSOME, *Orthomin, an iterative method for solving sparse sets of simultaneous linear equations*, in Proc. of the 4th Symposium on Reservoir Simulation, Society of Petroleum Engineers of AIME, 1976, pp. 149–159.
- [9] J. ERNEST WILKINS, JR., *A note on skewness and kurtosis*, Ann. Math. Statistics, 15 (1944), pp. 333–335.

